EUROGRAPHICS 2013 / B. Levy, X. Tong, and KK. Yin
(Guest Editors)

Volume 32 (2013), Number 7

Learning and Applying Color Styles From Feature Films

S. Xue[1,2], A. Agarwala[2], J. Dorsey[1], H. Rushmeier[1]

[1]Yale University
[2]Adobe Systems, Inc.

Figure 1: *The leftmost input photo is altered with color styles* horror, happy, David Fincher, *and* Wes Anderson *learned from feature films.*

Abstract

Directors employ a process called "color grading" to add color styles to feature films. Color grading is used for a number of reasons, such as accentuating a certain emotion or expressing the signature look of a director. We collect a database of feature film clips and label them with tags such as director, emotion, and genre. We then learn a model that maps from the low-level color and tone properties of film clips to the associated labels. This model allows us to examine a number of common hypotheses on the use of color to achieve goals, such as specific emotions. We also describe a method to apply our learned color styles to new images and videos. Along with our analysis of color grading techniques, we demonstrate a number of images and videos that are automatically filtered to resemble certain film styles.

1. Introduction

The colors and tones used in films are carefully choosen to maximize emotional impact. For example, filmmakers use warm colors to convey positive emotions, while high contrast and dark tones emphasize the bleakness of film noir plotlines. While perceptual scientists have long observed that colors are associated with specific emotions [Whe94, OLWW04a, OLWW04b], feature film directors and cinematographers are masters at exploiting this relationship. Directors use a process called "color grading" to manipulate color, and this process has received considerable attention in film studies; for example, Bellantoni [Bel05] hypothesizes correlations between various colors and the emotions of film, while Hurkman [Hur10] provides a practitioner's handbook of how to perform color grading with various emotional goals in mind.

However, to date the analysis of color grading techniques has been entirely qualitative, consisting of "rules of thumb" collected from practitioners, and observations from film studies experts. In this paper, we data-mine the work of the experts by analyzing the statistical properties of the use of color in feature films. Specifically, we collect a large set of feature film clips and manually attach labels, such as genre (e.g., romance), emotion (e.g., sad), director (e.g., David Fincher), and time period (e.g., 60s). We then learn a model that maps color and tone properties of the clips to the associated labels. This learned model allows a deeper, evidence-based understanding of the practice of color grading; we demonstrate the power of the model by testing existing hypotheses on color grading techniques, as well as proposing new ones. Our dataset and learned model also allow us to perform two more practical tasks. One, given a new, unknown film clip, we can predict various properties about it such as genre, emotion, or even the film's director. Two, we can transform an existing photograph or video to better exhibit one or more of the film properties we model. For ex-

ample, we can transform a photograph to look more like a horror film, or more like the films of Wes Anderson. While manually-designed photo filters exist and are widely-used (e.g., Instagram and Adobe SpeedGrade), our data-driven technique shows how we can automatically derive a large set of photo filters from the color grading techniques used by the masters of color manipulation in feature films.

2. Related Work

While most filmmaking books stress the importance of using color styles [Lan10, And12], few books offer specific rules to create them. Bellantoni [Bel05] discuss associations of common colors and emotions used in films; however, the relationships are multiple-to-multiple mappings that make it hard to design effective filters. Professional colorists [Hul08, Hur10] describe a three-step process for color grading: first, perform global luminance correction, second perform global adjustments of hue and saturation, and finally perform local adjustments on masked regions. Our pipeline also focuses on global color and tone manipulation with automatic local adjustments to protect memory colors.

We are not the first to analyze the low-level statistics of films. Brunick et al. [BCD13] analyze temporal trends of shot duration, brightness, and color in films over the past several decades. Others have used low-level statistics to predict the mood or genre of a film [WDC04, WC06, HX05]. These system typically use additional features such as motion and audio; we restrict ourselves to learning color styles, since they are most useful for image and video filters.

Creating image filters that achieve various emotions and moods is also a popular research topic. The systems in [MSMP11, YP08, WYW*10] typically use data from the existing color literature (e.g., [Whe94]) or color social media websites to create emotion or concept-labeled color themes, which are then transferred to images using color transfer techniques [RAGS01, PR10]. Csurka et al. [CSMS11] use similar data to learn a model mapping color themes to natural language concepts. In contrast, we learn our color styles from professional film clips, and do not use color themes as an intermediate representation.

Example-based methods are another approach to style enhancement. While color transfer techniques [RAGS01, KMHO09, PR10, HSGL11] exactly map the colors of a single example image to the target, Bonneel et al. [BSPP13] achieve temporally consistent color transfer between videos. Wang et al. [WYX11] learn models of style mapping between registered image pairs before and after adjustment. Our work build parametric models of styles from multiple examples, which allows more flexibility than a single example during stylization. Also, our method obviates the need for the user to find an example exhibiting a desired style.

Achanta et al. [AYK06] present a set of rules for manipulating videos to achieve various emotions; they manipulate properties such as framing and speed that go beyond our focus on color and tone. Also, our methods are derived directly from films rather than created by hand.

Our work shares the motivation of Doersch et al. [DSG*12], who use data to extract the visual signatures of urban architectural styles; we focus on characterizing the color styles of films. Finally, Palermo et al. [PHE12] analyze photographic color and tone in order to predict the decade in which a vintage photograph was taken; we similarly examine films to extract their color characteristics. However, they do not attempt to create filters.

Emotion	Genre	Director	Period	Location
happy	action	Tim Burton	60s	western
excited	comedy	David Fincher	70s	college
mysterious	crime	Peter Jackson	ancient	country
tender	drama	Wes Anderson		
neutral	fantasy	Coens Brothers		
melancholy	film-noir	Bong Joonho		
discomfort	history			
fear	horror			
sad	romance			
	war			

Table 1: *The labels of color styles from five categories.*

3. The Data-driven Study

We use a data-driven scheme to create models of color styles. We collect a dataset of 569 clips extracted from 52 mainstream feature films, each of which consists of one or several shots with consistent color styles. The durations of the clips range from 3 seconds to 5 minutes. To avoid handling DVDs and processing entire films, we extract the clips from film trailers and other short excerpts available online. Clips are temporally limited to single scenes so that the color style is continuous and constant. We select films that well sample our list of labels (e.g., we require several films of each of the directors that we study, several 60s and 70s films, etc.).

We manually label each clip by the styles it conveys. The style labels come from five categories: *Emotion*, *Genre*, *Director*, *Period*, and *Location*. The complete style labels are listed in Table 1. Multiple labels could be associated with one clip. For example, a single clip from David Fincher's film *Panic Room* could be simultaneously labeled with *fear*, *crime*, *film-noir*, and *David Fincher*.

Given the labeled training set, we first perform studies to identify the color styles that are visually distinctive and recognizable. We construct and analyze parametric models to describe these styles, and use these models to examine hypotheses on the practices of color grading.

Emotion					
happy	excited	mysterious	tender	neutral	melancholy
0.726	0.599	0.788	0.605	0.644	0.558
discomfort	fear	sad			
0.650	0.622	0.553			
Genre					
action	comedy	crime	drama	fantasy	film-noir
0.727	0.848	0.651	0.660	0.762	0.739
history	horror	romance	war		
0.740	0.711	0.679	0.593		
Director					
Tim Burton		David Fincher		Peter Jackson	
0.808		0.844		0.801	
Wes Anderson		Coen Brothers			
0.894		0.745			
Period			Location		
60s	70s	ancient	western	country	college
0.823	0.783	0.684	0.762	0.618	0.711

Table 2: *The classification accuracies* $G = \sqrt{T_{nr} \times T_{pr}}$ *of different style labels. A value of* 1 *indicates perfect classification.*

3.1. Identifying Distinctive Color Styles

We assume that a color style label is visually distinctive if it is easy to distinguish clips tagged with that label from clips that are not. We train supervised classifiers on each style of interest (positive samples have the style label, and negative samples do not); classifiers with high accuracy indicate distinctive styles.

We choose features that are most commonly used by professionals in color correction [Hur10]. The features come from three aspects of color: *luminance*, *hue*, and *saturation*. Since colorists often treat different bands of luminance separately, such as highlights, mid-tones, and shadows, we model luminance with a 10-bin histogram in the log2 domain to decode gamma. Colorists also treat hue in different luminance bands separately (e.g, by making the shadows redder). So, we compute a 10-bin histogram of hue in each of three bands (highlights, midtones, and shadows)[†]. For saturation, we compute the statistics within different luminance bands similarly; however, we use only the average of the highest saturation zone (top 0.1%) in log2 domain, as suggested by [XADR12]. We find that the highest saturation values most correspond to our perception of overall image saturation. In total, $10 + 10 \times 3 + 1 \times 3 = 43$ features are computed for each clip.

[†] Luminance bands are defined as highlights $(192, 255]$, midtones $(64, 192]$, and shadows $[0, 64]$ in gamma-encoded sRGB space.

We then train a binary classifier using AdaBoost, which builds a classifier as an ensemble of decision stumps. We experimented with a number of sets of features, and choose this set by trial-and-error. Too few features reduces accuracy while too many leads to overfitting. This set of features offers a reasonable trade-off between these two problems. We also experimented with nearest-neighbor and logistic regression, and found AdaBoost performed best. Also, we benefit from AdaBoost for interpreting selected features, as detailed in Section 3.2. Since our data is imbalanced (many more negative than positive samples for each label), we use a version of AdaBoost specialized for imbalanced data [SKWW07]. Specifically, during training the cost of a false positive is set to 1 while the cost of a false negative is set to c, where $c > 1$ increases the importance of positive samples. We set the value of c separately for each label using 10-fold cross-validation. We measure the performance of our classifiers in the presence of imbalanced training samples by combining the true negative rate T_{nr} and true positive rate T_{pr} into a single geometric mean $G = \sqrt{T_{pr} \times T_{nr}}$ [KHM98], again using 10-fold cross-validation. The results are listed in Table 2. We can observe that a subset of styles are very distinctive in terms of classification performance, while others are not.

3.2. Parametric Models of Color Styles

An AdaBoost classifier consists of a sequence of weak learners that each perform binary classification. The final AdaBoost output is a weighted sum $\mathcal{H}$ of the output values (1 or -1) of each weak learner; if $\mathcal{H} > 0$ the classification is positive and the clip conveys the color style. The magnitude of $\mathcal{H}$ is the margin of the prediction and can be interpreted as the *stylization strength* of an input clip. We use this strength value in our image filtering application in Section 4.

Each weak learner is a decision stump that simply compares a single feature against a threshold value. Values on one side of this threshold indicate that a clip is more likely to reflect the color style being tested. We can aggregate all these stumps into a visualization (Figure 2) indicating ranges of each feature that are preferred by a color style (brighter values indicate preferred ranges). To compute the values in this visualization, we first sum all the decision stumps. That is, for a decision stump with weight w and threshold k on a particular feature, we add $-w$ to all values below k, and w to all values above k in that feature's row in the diagram. Then, we shift and normalize all values so that the lowest value across all rows maps to 0 (black), and the highest value across all rows maps to 1 (white). This diagram visually shows the parametric model of color style represented by an AdaBoost classifier, and helps us to better understand complicated styles (Section 3.3.5). See supplemental materials for the visualization of other color styles.

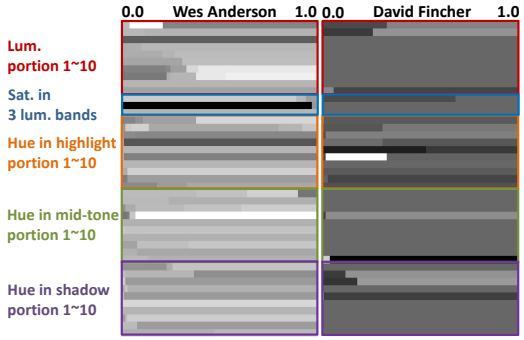

Figure 2: *The visualized AdaBoost classifier of the color style* Wes Anderson *and* David Fincher. *Each row represents a feature, which ranges from* 0.0 *to* 1.0 *after normalization. The preferred feature ranges for this style are brighter in intensities.*

3.3. Color Grading Hypotheses

Given our dataset and parametric model of color style, we can test existing hypotheses on color styles from the color grading literature, and formulate new hypotheses.

3.3.1. Positive Emotions

Bellantoni [Bel05] writes that warm colors are used to accentuate happy emotions in films. We thus test the null hypothesis that warmer colors have equal occurrence in *happy* clips and non-happy clips, against the alternative hypothesis that warmer colors occur more often in happy clips. We compute the portion of warm colors (hue in $[-30^\circ, 90^\circ]$, i.e., magenta, red, orange, and yellow) in every clip, and employ a one-tailed $t-$test to compare the average portions of warm colors in happy and non-happy clips. The rejection of N_0 against N_A is confirmed ($p^* = 3 \times 10^{-6}$), meaning that this simple rule is indeed used by filmmakers. Figure 3 shows the distribution of mean hues in happy and non-happy clips, which visually confirms our findings. See supplemental materials for more visualizations of the hypotheses we test in this section.

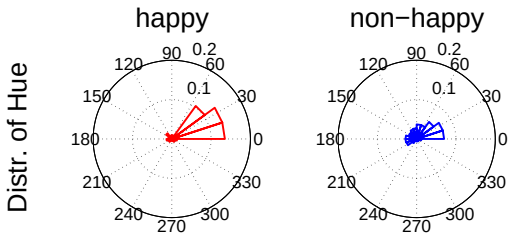

Figure 3: *The polar distributions of mean hues of happy clips and non-happy clips.*

Hurkman [Hur10] suggests boosting yellows in highlights, as well as blues in shadows to convey *tender* feelings. We compute the hue occurrence in highlight and shadows, finding that boosted yellow is indeed present ($p^* = 0.007$), while increased blues are not observed ($p^* = 0.99$).

Melancholy is one of the more complicated emotions in films. Hurkman suggests several techniques to achieve it: 1) a mild (brighter) black, 2) a pastel combination of greens and blues in the highlights, or 3) a pale blue cast in the shadows. Hypothesis 1 is quantitatively restated as the black level (lowest luminance) in melancholy clips are higher. However, the average blacks for melancholy are only slightly but not significantly higher (-10.93 vs -11.06 in log domain, $p^* = 0.29$). Second, we find blues to be slightly but not significantly stronger in highlights (occurrence 0.39 vs 0.35, $p^* = 0.124$), while green is very significantly boosted (occurrence 0.25 vs 0.17, $p^* = 0.0008$). Finally, checking the tones in shadows, we confirm a stronger blue cast (occurrence 0.26 against 0.20, $p^* = 0.0065$).

3.3.2. Negative Emotions

Both Hurkman and Bellantoni suggest using blue casts to convey *sad* and *fear* emotions. Our tests confirm this intuition. In sad clips, average occurrence of blue tones is 0.463, significantly higher ($p^* = 0.00033$) than 0.280 in non-sad clips. The fear clips show 0.4313 average blue occurrence while non-fear clips have only 0.2643.

The more subtle negative emotion of *discomfort* can be associated with multiple tones: red as defiant and anxious, orange as exotic and toxic, and green as corrupt and poisonous [Bel05]. Within our dataset, we do not observe a significant increase in red and orange in discomfort clips. This could be explained by the fact that these two colors are associated with many other emotions, both positive and negative.

On the other hand, we indeed find significantly more green in discomfort clips than in others, with occurrence 0.152 vs 0.091, and $p^* = 1.9 \times 10^{-5}$. Since green tones are less naturally used as an overall color cast, it has fewer associations with positive emotions and thus are widely used in discomfort styles.

3.3.3. Genres

Bellantoni proposes that red and yellow hues are used to heighten a sense of *romance*. We confirm this hypothesis: the average occurrence of red/yellow in romance flips is 0.676, much higher than 0.582 in non-romantic clips.

While [Bel05] describes the use of purple and blue in *fantasy* films, we do not find statistical evidence. Blue is slightly more common but purple show equal occurrence. However, we do find significantly more cyan in *fantasy* clips (with occurrence 0.325) than in non-fantasy clips (occurrence 0.167), with $p^* < 1 \times 10^{-7}$.

3.3.4. Vintage Styles

[Hur10] describes the following method to achieve a *vintage* style: 1) boost global contrast, 2) reduced saturation, and 3)

Figure 4: *The pipeline of manipulating an input frame. The insets show the decomposition of luminance bands, where yellow shows highlights, red shows midtones, and blue shows shadows. Hues/saturations in different bands are manipulated separately.*

add a yellow or magenta cast. We use our labels 60*s* and 70*s* as an approximation of vintage.

For global contrast, we do not find an average increase of global contrast in *60s* clips, but significantly increased contrast in *70s* clips. We do not observe a significant difference in saturation between *60s*/*70s* and other film clips. Lastly, higher occurrence of yellow is confirmed in both *60s* and *70s* clips, with p^* at 0.031 and 0.029, respectively. However, magenta has a slightly higher occurrence in *70s* clips but no increased occurrence in 60*s*.

3.3.5. Color Styles of Directors

In this section we examine the unique color styles of two directors, Wes Anderson and David Fincher. Note that we only examine low-level color statistics; both directors have many other aspects to their visual styles, from set designs to costume selection.

Figure 2 (left) visualizes the classifier associated with the *Wes Anderson* color style. By examining the darkest and brightest values, we can see that the prevalence of shadows is low (the white in the first row indicates a preference for a small value for the shadow bin), and the mid to high tones are generally large. Thus, the films of Wes Anderson tend to be bright with low global contrast and bright blacks. Numerically, his blacks are on average higher than other films (-10.12 versus -11.09 in log2 domain, with $p^* = 4.1 \times 10^{-5}$), average luminance is higher (-4.59 vs -6.01, $p^* = 2 \times 10^{-6}$), and highlights are less prevalent (-0.543 vs -0.331, $p^* = 0.0097$). For hue, we observe a prevalence of red-orange hues (63.1% vs 33.4% in other films, $p^* < 10^{-6}$). In supplemental materials we visualize the distributions of luminance and hue for Wes Anderson versus other films.

We next examine the color style of *David Fincher* using the patterns in Figure 2 (right). Luminance exhibits a strong pattern; shadows are very prevalent (first two rows in luminance), while highlights are rare (the 10th row in luminance). Overall, Fincher's films exhibit a low-key with large shadows. Numerically, the shadow portion is 24.1% versus 14.2% in other films ($p^* = 2 \times 10^{-6}$), and the highlight portion is 2.91% versus 5.27% in other films ($p^* = 0.0052$). For hue, we observe a high prevalence of green hues in the highlights (44.3% versus 25.0%, $p^* = 1 \times 10^{-6}$) and shadows (43.5% versus 16.8%, $p^* < 10^{-6}$). Green is not a typical lighting color, which suggests it is purposely added to create a suspenseful atmosphere. We show distributions of luminance and hue for Fincher's films in supplemental materials.

4. Adding Color Styles

In this section we describe a method to add one of our modeled color styles to a new film clip or still image. Existing mobile apps like Instagram as well color grading software like Adobe SpeedGrade offer many filters that can achieve a variety of color styles; however, they are all hand-coded. We show a methodology to automatically create filters from labeled data, and in particular, film clips labeled by style.

4.1. Overview

To add a color style $\mathcal{S}$ to an input clip v, we manipulate the same features of v used in our classifier in Section 3.1. Recall that these features were designed to model the same properties manipulated by professional color correctors. We use three categories of features: luminance (a 10-bin histogram), hue (three 10-bin histograms corresponding to hues in highlights, midtones, and shadows), and saturation (largest saturation bin in highlights, midtones, and shadows). Following the practice of colorists [Hul08, Hur10], we manipulate luminance, hue and saturation of the input clip in a sequential order. When manipulating hue, the hues in highlights, midtones, and shadows are edited sequentially (Figure 4).

When manipulating each feature F ($F \in \{$luminance, hue, saturation$\}$) of clip v, we first find k clips of style $\mathcal{S}$ from the training dataset (569 clips), creating a target set C. These clips are chosen according to two criteria: one, they should have similar features to clip v, and two, they should exhibit strong *stylization strength*. That is, they should be strong rather than subtle examples of the color style. Then, we perform a multiple-to-one style transfer from the clips in C to v. Specifically, we compute a target set of features by averaging the features across the clips in C. We then use the target features to update the original features of v.

4.2. Selected the target Set C

To select a target set C we first need to define a similarity measure d between features of two clips in terms of their features F and F'. The distance d is computed differently for luminance, hue, and saturation features. Note that we select a different set C of k clips for each feature; one set for luminance, three for hue (for highlights, midtones, and shadows), and three for saturation.

For luminance, we compute d as the L_1 distance over the 10-bin histogram, or $d(F,F') = \frac{1}{10}\sum_{i=1}^{10}|F_i - F'_i|$. For hue, there are three 10-bin histograms describing hues in highlights, midtones, and shadows, separately. We again use L_1 distance for each histogram, and select three sets of clips (one per histogram). Finally, the largest saturations bins in highlights, midtones, and shadows define the three scalar saturation features. We use Euclidean distance to select a different set C of clips for each scalar.

Selecting clips C similar to v prevents unnatural stylization of v. However, the clips C should also be strongly stylized to avoid overly subtle results. Recall that in Section 3.1 we compute a stylization strength $\mathcal{H}$ of a clip using its AdaBoost score. Here we use the part of $\mathcal{H}$ that is contributed only by feature F, defined as $\mathcal{H}_F$, to define the stylization strength of feature F. To select clips that are both similar and strongly stylized, we select the k clips that minimize

$$d(F,F') - s \cdot \mathcal{H}_{F'}, \tag{1}$$

where we set $s = 0.1$ and use $k = 3$ for all experiments in this paper.

4.3. Feature Manipulation

The next step is to compute a target feature F' by averaging the features from the selected clips C. F' is then used to update the original feature F of clip v, as follows.

Luminance F' is a new averaged 10-bin histogram; we use standard histogram matching [GW07] to convert the original luminance histogram F of v to F'.

Hue Since we have three source hue histograms corresponding to highlights, midtones, and shadows, separately, we apply histogram matching to each luminance band sequentially. However, if we simply use histogram transfer artifacts can result, as humans are very sensitive to odd hues. We therefore regularize the transfer in several ways, as described in Section 4.4.

Saturation Transfer from the original saturation histogram to a target saturation histogram is accomplished by simply shifting the original histogram to align the largest saturation bins, since we only use this bin as a feature. This shifting is done in the log2 domain, requiring the regularization techniques in Sections 4.4.3 and 4.4.4 to avoid artifacts.

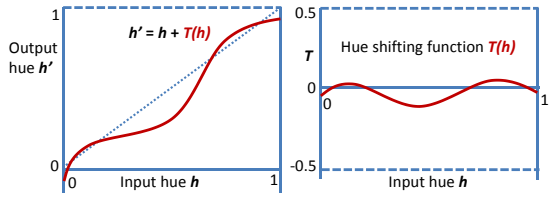

Figure 5: *The histogram matching process will map the original hue h to a new hue h', which is formulated by a a shifting function $\mathcal{T}(h) = h' - h$. Note that since h is a circular value, $\mathcal{T}(h)$ must satisfy circular continuity, i.e., $\mathcal{T}(0) = \mathcal{T}(1)$.*

4.4. Regularization on Hue Shifting Function

Histogram transfer maps each original hue value h to a new hue h'. We can define this transfer using a *hue shifting function* $\mathcal{T}(h)$, where $h' = h + \mathcal{T}(h)$, or $\mathcal{T}(h) = h' - h$ (Figure 5). $\mathcal{T}$ is a discrete function with input h uniformly discretized in $[0,1]$ with step $\triangle h = 0.001$. Note that we are dealing with hues in one luminance band; other bands are treated in the same manner. There are four problems with $\mathcal{T}$ that can cause artifacts. First, two originally close hues can be mapped to very different hues. Second, hues for objects with strong human expectations of specific colors, such as sky or skin, are mapped to unnatural hues. Third, hue computation can be unstable in certain areas, such as over-/under-exposed areas and low-saturation areas. Fourth, the hue transfer functions for three luminance bands are different. So, a visible jump may occur between two neighboring pixels of the same hue if their luminance values fall into two different bands. Rabin et al. [RDG10] use non-local filters to alleviate the first problem. Instead, we use the following regularization techniques to deal with all sources of artifacts.

4.4.1. Enforcing Continuity of Hue Shifting

Once an initial hue shifting function $\mathcal{T}$ is computed using histogram transfer, we smooth it by applying a 1D box filter, where the filter radius is 0.07. Since hue is a circular value defined in $[0,1]$, all computations about hue are applied in a circular manner.

A smoothed $\mathcal{T}$ can still map two originally close hues to very different hues if $\mathcal{T}$ lacks good *Lipschitz continuity*. That is, given two hues h_1 and h_2, shifted h_1^* and h_2^* should satisfy the Lipschitz condition $|h_1^* - h_2^*| \leq \beta|h_1 - h_2|$, where β is a constant. For hue transfer, we expect a relatively small β, which can be obtained by limiting the magnitude of the derivatives of $\mathcal{T}$ since large derivatives in $\mathcal{T}$ pull apart similar hues (Figure 6).

Therefore, we compute a better mapping function $\hat{\mathcal{T}}$ whose derivative $\hat{\mathcal{T}}'$ is the clipped version of $\mathcal{T}'$. We clip the derivatives within $[-\beta,\beta]$, where $\beta = 3$ in our experiments. Since we need to maintain the circular property of the hue

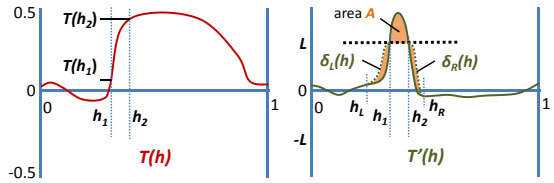

Figure 6: *We enforce better Lipschitz continuity of $\mathcal{T}$ by editing $\mathcal{T}'$. Note that $\mathcal{T}'$ also satisfies circular continuity, i.e., $\mathcal{T}'(0) = \mathcal{T}'(1)$.*

shifting function, i.e., $\hat{\mathcal{T}}(0) = \hat{\mathcal{T}}(1)$, we must guarantee

$$\int_0^1 \hat{\mathcal{T}}'(h)dh = \int_0^1 \mathcal{T}'(h)dh. \tag{2}$$

So, the clipped area A (where $A = \int_{h_1}^{h_2} \hat{\mathcal{T}}'(h)dh$, see the right in Figure 6) must be added back to $\hat{\mathcal{T}}'$. To minimize the effect on hues away far from clipped regions $[h_1, h_2]$, we boost the values of $\hat{\mathcal{T}}$ on both sides of $[h_1, h_2]$ so that $A/2$ is added to left and right sides. Formally, we compute two functions, δ_L and δ_R, to add to clipped $\mathcal{T}'$. Function δ_L is defined on $[h_L, h_1]$, and δ_R on $[h_2, h_R]$, and is otherwise 0. The delta functions must satisfy (Figure 6):

$$\int_{h_L}^{h_1} \delta_L(h)dh = \int_{h_2}^{h_R} \delta_R(h)dh = \frac{A}{2}. \tag{3}$$

We define $\delta_L(h)$ and $\delta_R(h)$ as

$$\delta_L(h) = \frac{h - h_L}{h_1 - h_L} \cdot (\beta - \mathcal{T}'(h)) \tag{4}$$

$$\delta_R(h) = \frac{h_R - h}{h_R - h_2} \cdot (\beta - \mathcal{T}'(h)) \tag{5}$$

where h_L and h_R are computed by plugging Eqn. 4 and 5 into Eqn. 3. If there are multiple peaks or valleys to be clipped, they are clipped sequentially.

Finally, we get the new derivative function $\hat{\mathcal{T}}'$, which is the sum of the clipped $\mathcal{T}'$ and δ_L and δ_R. We integrate back to obtain the new hue shifting function $\hat{\mathcal{T}}$, using the initial condition $\hat{\mathcal{T}}(h_0) = \mathcal{T}(h_0)$. h_0 is the farthest hue from the clipped region $[h_1, h_2]$. Note that since we use discretized $\mathcal{T}$ in practice, T' is computed by $\triangle T/\triangle h$, and integration is computed by summation.

4.4.2. Memory Color Protection

Humans have expectations for the colors of certain familiar objects, such as skin; these are often called *memory colors* [Bar60, YBDR99]. We protect the hues of two types of memory colors: skin and sky.

The average memory color of skin was first identified by [Bar60]. The Gaussian distribution of memory skin colors on screens was identified by a large-scale crowdsourced experiment [XMRD12]. Converted to *HSV* space, the hues of skin tones ranges approximately within $[0.0, 0.1]$. We then

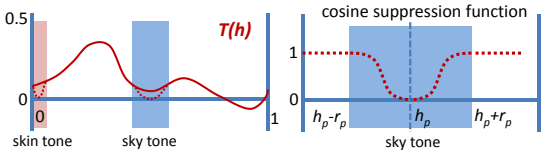

Figure 7: *On the left, we show the effect of memory color protection on $\mathcal{T}(h)$ (the dashed line is after protection). On the right, we show the cosine suppression function.*

Figure 8: *The results after hue manipulation with and without memory color protection.*

add a protection to hues within this range. See Figure 7 for illustration. Let the hue shifting function be $\mathcal{T}(h)$; for the hues within $[0.0, 0.1]$, we modify the function as

$$\tilde{\mathcal{T}}(h) = \frac{1}{2}\left(1 - cos(\frac{d(h,h_p)\pi}{r_p})\right)\mathcal{T}(h) \tag{6}$$

where $d(h,h_p)$ is the distance between h and the center hue $h_p = 0.05$ within the skin tone range, and the $r_p = 0.08$ is the radius of this range. We can see in Figure 7 that if h is outside the skin tone range, $\mathcal{T}(h)$ is not affected at all. The closer h gets in the skin tone range, the more $\mathcal{T}(h)$ is suppressed. We apply similar protection to sky hues on the shifting function $\mathcal{T}(h)$, where the protection range is defined in $[0.45, 0.71]$. Figure 8 shows an example.

4.4.3. Unstable Hue Computation

Since hue computation is unstable and inaccurate for pixels in over-exposed, under-exposed, and low-saturation pixels, we scale down the magnitudes of hue shifting on these pixels. Notice that this regularization method varies spatially across the image. The under-exposed and low-saturation pixels have low chroma, where chroma is the product of saturation and brightness values. When applying a hue shifting function $\mathcal{T}$ to a pixel p with chroma $C(p) < t_C$, where $t_C = 0.1$ is a preset threshold, we modify $\mathcal{T}$ by

$$\tilde{\mathcal{T}}(h) = \frac{1}{2}\left(1 - cos(\frac{C(p)\pi}{t_C})\right)\mathcal{T}(h). \tag{7}$$

Similarly, we set a threshold $t_\ell = 0.58$ [‡] to scale down $\mathcal{T}$ on over-exposed pixels,

$$\tilde{\mathcal{T}}(h) = \frac{1}{2}\left(1 - cos(\frac{1-\ell(p)\pi}{1-t_\ell})\right)\mathcal{T}(h) \tag{8}$$

[‡] Equal to luminance 200/255 in gamma-encoded sRGB.

Figure 9: *Left: the input image. Middle:* Wes Anderson *style added. Artifacts due to hue shifting in over-/under-exposed pixels. Right: toning down the shifting for unstable hues significantly alleviates the artifacts.*

where $\ell(p)$ is the luminance of gamma-decoded sRGB pixels. See Figure 9 for an example. We apply the same regularization to the saturation shift function.

4.4.4. Transition Between Luminance Bands

We compute three different hue shifting functions separately for highlights, midtones, and shadows. If two neighboring pixels have the same hue, but their luminance values fall into different bands, the transfer functions may yield very different hues. To fix this issue, we linearly cross-fade the three shifting functions across the two transition regions between the three bands.

We define the boundaries of the three luminance bands in Section 3.1. The transition region is defined with a symmetric margin of 40 (out of 255) in gamma-encoded sRGB space. When cross-fading two hue shifting functions, circularity is enforced. The cross-fading between luminance bands are also employed for the saturation shifting function, where the margin of transition zone is instead set to 20.

4.5. Results and Validation

We demonstrate our technique for adding color styles by applying them to a variety of videos and photos. Processing each frame takes less than 5 seconds in our experiments. Figure 10 shows the results of adding four different styles to a number of examples; see supplemental materials for more. The fifth, sixth, and seventh examples in Figure 10 are applied to video clips (though we show single frames in the paper), while the others are still photos. See supplemental materials for more photos and videos. It is important to note that these styles have a number of characteristics that we do not model; we only model low-level color and tone statistics. Since content can easily overwhelm style, we test on images that are as neutral as possible.

We evaluate the effectiveness of adding color styles by performing a human subjects study. Each individual test asks a user to compare a filtered image with its original, and then asks which version better reflects a color style. The order of image is randomized. We test two emotions (happy and mysterious), two genres (horror and comedy), and two directors (David Fincher and Wes Anderson), which are identified as distinctive styles. We use Amazon Mechanical Turk to test the emotions and genres. Since familiarity with director style is much less universal, we use a study of our peers who claim knowledge of the two directors to evaluate these styles. The methodology of experiments are the same for the MTurk study and the peer survey. Each subject is presented with 24 pairs of actual test images (still photos and frames of short clips, see supplemental materials); one is the original and the other with one random style added. Three test pairs with obvious answers are added to test the quality of the test subjects. Users who fail one or more test pairs are removed from analysis. We collected 178 responses from MTurk, and removed 53 users. We collected 24 peer responses and removed 6 users. We perform a one tailed t-test to test the null hypothesis N_0 (the votes for the edited and original images are equal) against the alternative hypothesis N_A (the votes for the edited image are significantly higher). Table 3 show the results across all images. See supplemental materials for the t-test results on individual images. Also, we include the results of adding color styles on longer clips in the supplemental materials, which demonstrate the temporal consistence of the color altered clips. The stability is guaranteed by the fact that a mapping function is learned and then equally applied throughout all frames.

Style	*horror*	*happy*	*comedy*	*mysterious*
p^*	$< 10^{-7}$	0.3912	0.3325	$< 10^{-7}$
Style	*David Fincher*	*Wes Anderson*		
p^*	$< 10^{-7}$	2.4×10^{-5}		

Table 3: *The t-test p^* values of N_0 against N_1 across all images. The smaller p^* is, the more effective the edited images are in terms of representing color styles.*

We can see that our method effectively models the horror genre, mysterious emotion, and the two directors. Our method is less successful with happy and comedy. This more poor result is likely because these two filters are fairly subtle, and have a fairly standard appearance that is already fairly consistent with most neutral images. This reveals the clear limitation of data-driven modeling of color styles that it is more effective for strong and visually distinctive styles than more normal ones.

5. Limitations and Future Work

Our study has a number of limitations. First, our database of film clips is certainly not exhaustive, so our observations are limited by dataset bias. We tried to collect a large enough sample for the specific labels that we modeled, but we cannot be certain that our conclusions will not change with more samples. However, our methodology should be applicable to larger datasets. Also, our style transfer method mostly targets shifts in overall color and tone; many successful com-

mercial filters also employ other methods like vignetting or image degradation. Finally, our method uses a number of thresholds and parameters. Ideally, we would expose a small number of them so that users can fine-tune their results and trade-off between fidelity to the original image and the strength of the style.

As future work we would like to explore searching of online video and photo databases, so that users can search for clips with certain styles that work well together. We would also like to model styles of other media; for example, modeling styles of vector illustration would help users to build style-consistent illustrations from clip-art components.

References

[And12] ANDERSSON B.: *The DSLR Filmmaker's Handbook: Real-World Production Techniques.* Sybex Press, 2012. 2

[AYK06] ACHANTA R., YAN W.-Q., KANKANHALLI M. S.: Modeling intent for home video repurposing. *IEEE MultiMedia 13*, 1 (Jan. 2006), 46–55. 2

[Bar60] BARTLESON C. J.: Memory colors of familiar objects. *J. Opt. Soc. Am. 50*, 1 (Jan 1960), 73–77. 7

[BCD13] BRUNICK K. L., CUTTING J. E., DELONG J. E.: Low-level features of film: What they are and why we would be lost without them. *Psychocinematics* (Jan. 2013), 133–148. 2

[Bel05] BELLANTONI P.: *If It's Purple, Someone's Gonna Die: The Power of Color in Visual Storytelling.* Focal Press, 2005. 1, 2, 4

[BSPP13] BONNEEL N., SUNKAVALLI K., PARIS S., PFISTER H.: Example-based video color grading. *ACM Trans. Graph. 32*, 4 (July 2013), 39:1–39:12. 2

[CSMS11] CSURKA G., SKAFF S., MARCHESOTTI L., SAUNDERS C.: Building look & feel concept models from color combinations: With applications in image classification, retrieval, and color transfer. *Vis. Comput. 27*, 12 (Dec. 2011), 1039–1053. 2

[DSG*12] DOERSCH C., SINGH S., GUPTA A., SIVIC J., EFROS A. A.: What makes paris look like paris? *ACM Trans. Graph. 31*, 4 (July 2012), 101:1–101:9. 2

[GW07] GONZALEZ R. C., WOODS R. E.: *Digital Image Processing (3rd Ed.).* Prentice Hall, 2007. 6

[HSGL11] HACOHEN Y., SHECHTMAN E., GOLDMAN D. B., LISCHINSKI D.: Non-rigid dense correspondence with applications for image enhancement. *ACM Trans. Graph. 30*, 4 (July 2011), 70:1–70:10. 2

[Hul08] HULLFISH S.: *The Art and Technique of Digital Color Correction.* Focal Press, 2008. 2, 5

[Hur10] HURKMAN A. V.: *Color Correction Handbook: Professional Techniques for Video and Cinema.* Peachpit Press, 2010. 1, 2, 3, 4, 5

[HX05] HANJALIC A., XU L.: Affective video content representation and modeling. *Multimedia, IEEE Trans. 7*, 1 (feb. 2005), 143 – 154. 2

[KHM98] KUBAT M., HOLTE R., MATWIN S.: Machine learning for the detection of oil spills in satellite radar images. *Machine Learning 30* (1998), 195–215. 3

[KMHO09] KAGARLITSKY S., MOSES Y., HEL-OR Y.: Piecewise-consistent color mappings of images acquired under various conditions. In *ICCV* (2009), pp. 2311–2318. 2

[Lan10] LANCASTER K.: *DSLR Cinema: Crafting the Film Look with Video.* Focal Press, 2010. 2

[MSMP11] MURRAY N., SKAFF S., MARCHESOTTI L., PERRONNIN F.: Towards automatic concept transfer. In *Proceedings of NPAR* (2011), pp. 167–176. 2

[OLWW04a] OU L.-C., LUO M. R., WOODCOCK A., WRIGHT A.: A study of colour emotion and colour preference. part i: Colour emotions for single colours. *Color Research and Application 29*, 3 (2004), 232–240. 1

[OLWW04b] OU L.-C., LUO M. R., WOODCOCK A., WRIGHT A.: A study of colour emotion and colour preference. part ii: Colour emotions for two-colour combinations. *Color Research and Application 29*, 4 (2004), 292–298. 1

[PHE12] PALERMO F., HAYS J., EFROS A. A.: Dating historical color images. In *European Conference on Computer Vision* (2012), pp. 499–512. 2

[PR10] POULI T., REINHARD E.: Progressive histogram reshaping for creative color transfer and tone reproduction. In *Proceedings of NPAR* (2010), pp. 81–90. 2

[RAGS01] REINHARD E., ASHIKHMIN M., GOOCH B., SHIRLEY P. S.: Color transfer between images. *IEEE Computer Graphics & Applications 21*, 5 (Sept./Oct. 2001), 34–41. 2

[RDG10] RABIN J., DELON J., GOUSSEAU Y.: Regularization of transportation maps for color and contrast transfer. In *Image Processing IEEE Intl. Conf. on* (2010), pp. 1933–1936. 6

[SKWW07] SUN Y., KAMEL M. S., WONG A. K. C., WANG Y.: Cost-sensitive boosting for classification of imbalanced data. *Pattern Recogn. 40*, 12 (Dec. 2007), 3358–3378. 3

[WC06] WANG H. L., CHEONG L.-F.: Affective understanding in film. *Circuits and Systems for Video Technology, IEEE Trans. 16*, 6 (june 2006), 689 – 704. 2

[WDC04] WEI C.-Y., DIMITROVA N., CHANG S.-F.: Color-mood analysis of films based on syntactic and psychological models. In *Multimedia and Expo, IEEE Int. Conf.* (june 2004), vol. 2, pp. 831 –834. 2

[Whe94] WHELAN B.: *Color Harmony 2*. Rockport Publishers, 1994. 1, 2

[WYW*10] WANG B., YU Y., WONG T.-T., CHEN C., XU Y.-Q.: Data-driven image color theme enhancement. *ACM Trans. Graph. 29*, 6 (Dec. 2010), 146:1–146:10. 2

[WYX11] WANG B., YU Y., XU Y.-Q.: Example-based image color and tone style enhancement. *ACM Trans. Graph. 30*, 4 (July 2011). 2

[XADR12] XUE S., AGARWALA A., DORSEY J., RUSHMEIER H.: Understanding and improving the realism of image composites. *ACM Trans. Graph. 31*, 4 (July 2012), 84:1–84:10. 3

[XMRD12] XUE S., MCNAMARA A., RUSHMEIER H., DORSEY J.: Crowd sourcing memory colors for image enhancement. In *ACM SIGGRAPH 2012 Talks* (2012), p. 48:1. 7

[YBDR99] YENDRIKHOVSKIJ S. N., BLOMMAERT F. J. J., DE RIDDER H.: Color reproduction and the naturalness constraint. *Color Research Application 24*, 1 (1999), 52–67. 7

[YP08] YANG C.-K., PENG L.-K.: Automatic mood-transferring between color images. *Computer Graphics and Applications, IEEE 28*, 2 (March-April 2008), 52 –61. 2

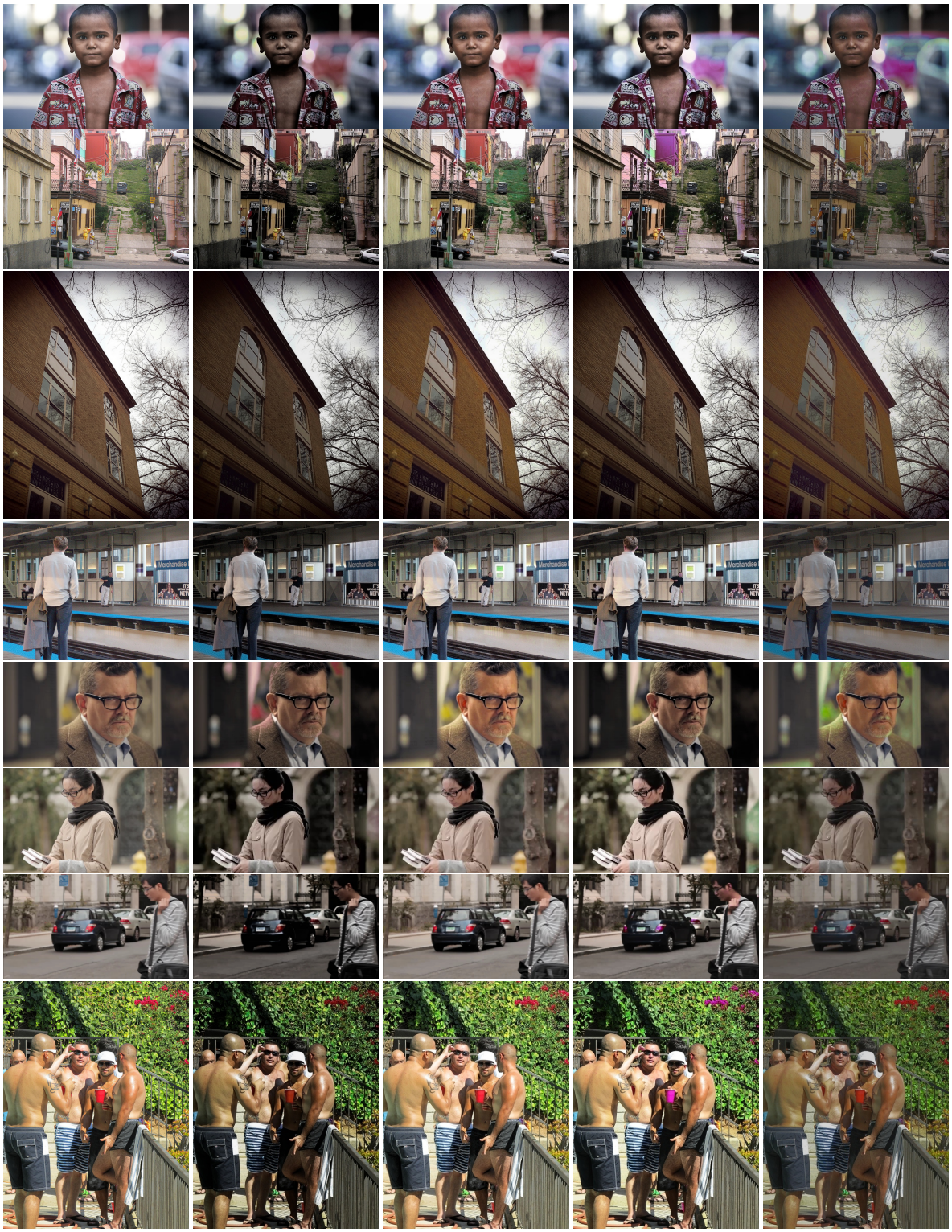

Figure 10: *The input images (leftmost) are altered with color styles (from left to right):* horror, happy, David Fincher, *and* Wes Anderson *learned from feature films. See supplemental materials for more results.*